

Published in final edited form as:

Res Soc Work Pract. 2004 May ; 14(3): 191–200. doi:10.1177/1049731503257882.

Cognitive Pretesting and the Developmental Validity of Child Self-Report Instruments: Theory and Applications

Michael E. Woolley, Gary L. Bowen, and Natasha K. Bowen

School of Social Work, The University of North Carolina at Chapel Hill

Abstract

Objective—In the context of the importance of valid self-report measures to research and evidence-based practice in social work, an argument-based approach to validity is presented and the concept of developmental validity introduced. Cognitive development theories are applied to the self-report process of children and cognitive pretesting is reviewed as a methodology to advance the validity of self-report instruments for children. An application of cognitive pretesting is presented in the development of the Elementary School Success Profile.

Method—Two phases of cognitive pretesting were completed to gather data about how children read, interpret and answer self-report items.

Results—Cognitive pretesting procedures identified validity problems with numerous items leading to modifications.

Conclusions—Cognitive pretesting framed by an argument-based approach to validity holds significant potential to improve the developmental validity of child self-report instruments.

Valid measurement is a fundamental component of any scientific endeavor. In social work practice and research we frequently rely on client self-report instruments to measure latent variables. These measures can be central to the development, evaluation, implementation, and dissemination of evidence-based social work practice (Gambrill, 1999). To support the growing emphasis on evidence-based practice, it is essential for social workers to develop and validate client self-report measures that target the variables important to practice from a social work perspective. This is especially true when it comes to instruments for children. Building on child development theory and emerging cognitive methods, we present *cognitive pretesting* as a scale development method that social work researchers can use to advance the validity of self-report instruments for children.

We begin with a discussion of possible obstacles to valid measurement with children. We then introduce the concept of *developmental validity*. To ascertain the developmental validity of items intended for child respondents, we must assess whether children can read, comprehend and answer the questions we believe we are asking. Therefore, we will discuss cognitive development theory to illuminate the cognitive demands children in middle childhood encounter while reading and answering self-report instrument items.

Correspondence concerning this article should be addressed to Michael E. Woolley, School of Social Work, The University of North Carolina at Chapel Hill, 301 Pittsboro Street Campus Box 3550, Chapel Hill, North Carolina, 27599-3550; mwoolley@email.unc.edu.

An earlier version of this article was presented at the 32nd Annual Theory Construction and Research Methodology Workshop (TCRM), Houston, Texas, November 2002. The authors would like to thank our TCRM discussants, Alan C. Acocock and Mark J. Benson, for their valuable feedback. The ESSP was developed in collaboration with Flying Bridge Technologies with funding from the National Institutes of Health, National Institute on Drug Abuse (IRYZDA13865-01, 3RYIDA13865-0151, AND 2 RY2DA013865-02).

Cognitive pretesting (DeMaio & Rothgeb, 1996; Foddy, 1998; Jobe & Mingay, 1989) is discussed as a qualitative methodology that assesses developmental validity. After describing the general methodology, we present a specific example of the application of cognitive pretesting during the development of a new instrument for third to fifth graders—the Child Form of the Elementary School Success Profile (ESSP) (N. K. Bowen, Bowen, & Woolley, in press). Finally, we present findings related to the developmental validity of items on the ESSP as well as lessons learned during the application of cognitive pretesting procedures with children.

Validity: An Argument-Based Approach

Social workers endeavor to measure the social reality of individuals using methods of self-report. Data collected by these methods are used for numerous purposes, from testing theories about human development and behavior to planning intervention and prevention activities. Frequently, self-report instruments consist of one or more sets of multiple-choice items, which are commonly referenced as scales. Each set of items is specifically designed to measure an underlying latent construct (DeVellis, 1991). Because social workers often base what they know and do on data gathered with self-report instruments, the validity of such instruments can have a profound impact on the nature of research findings and practice activities.

An instrument is said to be valid when it measures what it purports to measure (Thyer, 2001). This widely accepted definition of validity views it as a characteristic of an item, scale or instrument. Kane (1992) has presented an alternative perspective, an argument-based approach to validity. In this approach, validity is a characteristic of the *interpretation* ascribed to a score on an instrument. The interpretation of a score is built on the series of inferences leading to the generation of the score, from the design of the instrument to its administration. This chain of inferences constitutes the argument for the validity of any single score and typically includes: a theory base, assumptions based on that theory, empirically/statistically grounded assertions, clinical/observational knowledge, and commonly accepted procedures and methods. Kane (1992) asserts the most serious threats to validity are the weakest links in this chain of inferences—questionable or poorly supported assumptions. We propose the weakest link in the accepted methods used to design instruments for children is the assumption that children will interpret the items and answer options the way the adult instrument designers intended. Our research presents a systematic approach to using children as the best informants when pretesting the validity of self-report items intended for use with children.

Developmental Validity

The term developmental validity has been applied previously in the context of the development of social interventions (Thomas, 1985). In the context of designing self-report instruments for children, we apply the term developmental validity to describe when an item can be read, comprehended, and validly responded to by children in a targeted age range. To determine developmental validity we need to examine the relation between a self-report item (stimulus) and the child's answer (response).

This examination starts with assessing whether children can read items. Next, it looks at how children interpret items in terms of cognitively constructing an understanding of the words, sentences, concepts, and the links between those concepts. Then it is critical to assess whether children give valid responses, given the intent of the question and the explanation given by children for each chosen response. Establishing the developmental validity of an item strengthens the critical inferences—validity chain links—that children can read, interpret, and answer the item as the instrument designers intended.

Past research has shown that context and item wording can cause adults to misinterpret items and give invalid responses (Schwarz, 1999). Compared to adults or adolescents, children have more limited reading skills, vocabulary, attention span, and cognitive capacity to mentally represent and manipulate constructs both concrete and abstract. This would indicate that the cognitive processing of self-report items by children should be a central issue to designers of instruments for children.

Many child self-report instruments are in use, including instruments to measure depression (Wright-Strawdermann & Watson, 1992); anxiety (March, Parker, Sullivan, Stallings, & Connors, 1997); risk behaviors (Tinsley & Holtgrave, 1997); self-concept (McGuire et al., 1999; Montgomery, 1994); behavioral and emotional disorders (McConaughy, 1993); abuse and neglect (Amaya-Jackson, Socolar, Hunter, Runyan, & Colindres, 2000); guilt and shame (Ferguson, Stegge, Eyre, Vollmer, & Ashbaker, 2000); friendship (Yugar & Shapiro, 2001); posttraumatic symptoms (Greenwald & Rubin, 1999); and community violence (Cooley, Turner, & Beidel, 1995). Such measures are widely used in research and practice, however, reports on the development and quality of these measures do not include systematic evaluations of the way children read, comprehend and respond to individual items.

Concerns may be raised about the validity of existing instruments administered to children when the reading level of those instruments is well above the reading level of the intended population. For instance, we assessed the reading level of items from a number of child self-report measures for maltreatment that were reviewed by Amaya-Jackson et al. (2000). Flesch-Kincaid grade level reading scores on some instruments were well above the age of the target populations. Examples included an instrument for use with 11 to 17 year-olds (approximately grades 6 to 12) that had item reading levels as high as grade 9.0; an instrument for use with 5 to 20 year-olds (approximately grades kindergarten to college) that included item reading levels as high as grade 5.9; and an instrument for use with 10 to 17 year-olds (approximately grades 5 to 12) with item reading levels as high as grade 10.7.

Similarly, validity concerns arise when instruments designed for one age group are administered to a younger age group. One example is the use of an instrument designed for adolescents as a research tool in a study involving children after simply substituting the word “kid” for “teenager” in the instrument (McGuire et al., 1999). Finally, validity concerns arise when there is more than one source of data about the child. Child self-report data are often supplemented with data gathered from significant adults who know the child. In studies where both the child and a parent/guardian or teacher completed an instrument measuring constructs about the child, there is often very low concordance between the child’s and the adult’s responses (March et al., 1997; Montgomery, 1994; Tinsley & Holtgrave, 1997; Wright-Strawdermann & Watson, 1992; Yugar & Shapiro, 2001). While this is sometimes interpreted as indicating children cannot provide valid self-report data, it may also indicate children formulate different interpretations of items designed to measure the same underlying constructs.

An examination of such studies raises some serious and perplexing questions. For instance, what is a child’s capacity to validly self-report social reality? Why would self-report data gathered from children differ from the reports of adults who best know the child? How much of that discrepancy is due to systematic measurement error, and how much is due to true differences in perceived social reality? Are the child’s self-report measures or the significant adults’ measures more valid? If the child is the subject of research or intervention, is his or her perception more important than a significant adult’s perception? Although these questions offer an ambitious agenda for future research, the present discussion focuses on the question of a child’s capacity to self-report validly about his or her social reality. The

aim is to design child self-report items that are as valid as possible, given the still-developing cognitive capabilities of the target population. We now turn to child development theories to inform us about the capabilities of children to self-report their social reality.

Child Development Theories

Reading and answering a self-report instrument requires multiple simultaneous cognitive processes. Children must read and interpret the item, hold the concepts in their mind, search their memory for relevant information, read and interpret the response options, evaluate the item in the context of the response options, and then choose the option that best represents their answer. According to both Piagetian and information processing developmental theories, children possess more limited cognitive ability than adults to perform all of these tasks (Goldhaber, 2000; Goswami, 2002). However, those same theories also suggest that children are capable of validly responding to items if the wording and format fall within their cognitive capacity.

Piagetian Cognitive Development Theory

Our work focuses on children in Piaget's concrete operations stage of cognitive development (approximately six to twelve years of age). The concrete operations stage begins when a child can mentally represent things going on in the environment, can mentally manipulate those representations, and can draw conclusions about those manipulations. Concrete operations seems to describe the minimal cognitive skills necessary to read, comprehend, and respond to a self-report item. Therefore, Piagetian theory predicts children in the concrete operations stage are capable of giving valid self-report data about their social environment.

Nonetheless, the cognitive capacity of children during the concrete operations stage has developmental boundaries, and it is reasonable to anticipate that the success of specific items would be a function of the cognitive demands of those items. Therefore, child self-report items should be evaluated to ensure their cognitive demands do not exceed the cognitive capacity of the intended age range. Although Piagetian theory offers some developmental boundaries for anticipating the cognitive abilities of children in different age groups, the theory is limited in its usefulness to model the actual cognitive processes behind self-report. Information processing theory, our next focus, provides a means to examine those cognitive processes.

Information Processing Theory

In the late 1980s, researchers started applying the increasingly popular theory of information processing to survey design. Applying an information processing perspective, Hastie (1987) postulated that the fundamental component of analysis is information, which in the case of self-report instruments comprises words, phrases, concepts, and the items formed by these elements. Hastie also discussed the structure of memory from a social information processing perspective, examining how concepts are linked as nodes within lists of related concepts. A social information processing theory of memory seems applicable to self-report item response because responding to an item requires the respondent to access encoded memory structures. Hastie pointed to information processing as a promising theory in terms of survey methodology but did not go beyond a theoretical discussion of its possible implications.

Nearly ten years later Sudman, Bradburn, and Schwarz (1996) employed an information processing perspective to portray the cognitive process of responding to self-report items. According to their model, such a process starts with an interpretation of the item; the critical

issue here is whether the respondent interprets the item as intended. Next, the respondent determines if he or she holds a previously formed opinion. If yes, then he or she can directly formulate and give a response. If not, the respondent must access relevant information from memory, process that information within the context of the item, “compute” a response, and finally format and report that response. This model seems parsimonious and utilitarian, and although the assessment strategy was formulated for adults, it informs the cognitive pretesting model presented later in this article.

Cognitive Pretesting

In 1983 the National Center for Health Statistics organized the Seminar on the Cognitive Aspects of Survey Methodology, which brought together a group of cognitive psychologists and survey methodologists (Jabine, Straf, Tanur, & Tourangeau, 1984). The goal of that meeting was to develop a methodology to increase the validity of self-report survey questions through the application of theories of human cognition. Since that meeting there has evolved “considerable agreement that a better understanding of the cognitive processes in producing survey information . . . will result in better data” (Alwin, 2001, p. 19). There has also been increased recognition by instrument designers that “question testing should include procedures routinely used by cognitive psychologists to learn what respondents are thinking when they are trying to answer questions” (Fowler, 1995, p. 110).

The literature reports various *cognitive methods* for assessing the cognitive processing of self-report items by adult respondents. Examples include: (a) asking respondents to think aloud as they read, consider, and respond to items; (b) asking respondents to report what went through their minds as they responded to the item; (c) asking respondents to identify defects in the items; and (d) asking respondents to rate the comprehensibility of the items (Foddy, 1998; Jobe & Mingay, 1989; Sudman et al., 1996; Tourangeau, Rips, & Rasinski, 2000).

Although cognitive methods appear in the literature since the 1980s (Jabine et al., 1984; Jobe & Mingay, 1989), the techniques are underutilized (Thyer, 2001) and have received little empirical evaluation. One exception is a study by Foddy (1998) that applied several different cognitive methods to the same eight self-report items in order to evaluate which procedures most consistently identified validity problems. Foddy concluded that traditional pilot or field testing of questions was clearly inadequate to identify problems with items and that several of the cognitive methods proved effective. However, Foddy points out procedures utilized are seldom clearly described and vary from study to study. Foddy also asserts researchers have yet to demonstrate empirically that modifications made to items as the result of cognitive pretesting improve the accuracy of respondent interpretation or the resultant responses. Other authors have also called for more systematic and replicable procedures and empirical evaluations of cognitive methods (Tanur, 1999; Willis, DeMaio, & Harris-Kojetin, 1999).

While various cognitive methods described in the literature are reported to be valuable procedures and may be effective with adults, they may not be structured adequately for obtaining the necessary information from children. Some cognitive methods, however, do seem potentially appropriate for use with children. For instance, one method asks respondents to repeat items in their own words or to define key words in the item. Another method involves the interviewer repeating the respondent’s answer and then asking an open-ended question intended to determine how the respondent chose that answer, or asking the respondent to elaborate on the chosen answer (DeMaio & Rothgeb, 1996; Foddy, 1998). The cognitive method of interviewing a respondent while he or she responds to an item is

variously termed pretesting, cognitive testing, or cognitive interviewing; we will refer to it as *cognitive pretesting*.

Essentially, cognitive pretesting as we define it here involves conducting a structured interview with an individual as he or she reads, interprets, and answers an instrument item. The goals of the method are to assess four steps in the self-report process: (a) comprehension (interpreting the item accurately); (b) retrieval (adopting the appropriate perspective for the item); (c) judgment (understanding the response continuum and the response options within the context of the item); and (d) response (providing an answer and demonstrating an ability to provide a rationale for the answer) (DeMaio & Rothgeb, 1996; Jobe & Mingay, 1989; Tourangeau, Rips, & Rasinski, 2000). It should be noted that this process, which was conceptualized for adults, assumes that respondents will be able to read the item, an assumption that should not be made with children. Therefore, our proposed cognitive pretesting model for children (Figure 1) begins with the respondent reading the item, and then reflects the self-report process described above.

The need to accurately assess cognitive processing of items and scales is arguably even more important in the case of children because of cognitive developmental issues. A literature search, however, identified only one article applying cognitive pretesting procedures to instrument development with children (Rebok et al., 2001). While the extent to which scale developers may conduct other procedures labeled “pretesting” is not clear, it is clear that no systematic results or descriptions of the procedures described in this article are routinely presented in articles or sourcebooks on available self-report instruments for children.

Applying Cognitive Pretesting Procedures: The Elementary School Success Profile

The Elementary School Success Profile (ESSP), a drug abuse prevention screening tool for third to fifth grade children (approximately ages 7 to 11), is currently under development at the School of Social Work, The University of North Carolina at Chapel Hill. Development of the ESSP is sponsored by a grant from the National Institute on Drug Abuse (N.K. Bowen, Bowen, & Woolley, in press). This instrument for elementary school children evolved from an existing middle- and high-school instrument, the School Success Profile (SSP) (G. L. Bowen & Richman, 1993, 1997, 2001). Like the SSP, the ESSP measures ecological variables within a child’s central microsystems, including neighborhood, school, family, peer group, and health and well-being. Within each of these microsystems the ESSP measures both risk and protective factors that have been shown to influence child developmental outcomes. These measurements yield results that are useful for informing, monitoring, and evaluating school-based intervention and prevention activities focused on students at risk of school failure (G. L. Bowen, Woolley, Richman, & Bowen, 2001). The reliability and validity of scales on the SSP have been established over a decade of use with tens of thousands of middle and high school students.

Designing a Developmentally Valid Instrument

Our goal was to make the ESSP developmentally valid, and this goal influenced all stages and aspects of the design process. The first task in the development of the ESSP was to group items and scales on the SSP into three categories: (a) those items best answered by the child, (b) those items best answered by a parent/guardian, and (c) those items best answered by the teacher. This important step in the creation of a valid data-collection instrument for use with children acknowledged that while the inclusion of information from children is critical to comprehensive assessment, children are not the best sources of information on

some topics—such as the behaviors of adults in their neighborhood or their own school performance. Therefore, the ESSP includes a Child Form, a Parent Form, and a Teacher Form, and data from all three forms are reported on the individual profile generated for each child. The ESSP Child Form (ESSP-CF) is a computerized and animated computer program; the parent and teacher forms are paper and pencil surveys. Cognitive pretesting can be applied to any instrument for children no matter the format. Because the issue of developmental validity is most salient to the child form of the ESSP, the remainder of this article focuses on the development of the ESSP-CF. Information about the parent and teacher forms is available from the authors.

Wording and answer options of SSP items chosen for the ESSP-CF were modified given the lower reading, vocabulary, and comprehension levels of third to fifth graders. The ESSP research team performed this task, employing their expertise in areas of child development, education, and clinical practice with children. New items and scales relevant for children in middle childhood were also developed. The initial draft of the child form was then sent to five experts for consultation regarding developmental appropriateness. These consultants included two child development experts, a child and adolescent psychiatrist with expertise in child assessment, a scale development expert, and a professor of education with expertise in child development and assessment. The consultants provided valuable feedback about item wording, answer options, and formatting, resulting in modifications to the child form.

After developing the preliminary computerized form containing items we believed were worded appropriately for children aged 7 to 11, we sent the form back to the five expert consultants and again incorporated their input into the form. We also obtained and integrated feedback from nine teachers regarding the computerized format's appropriateness. At this point in the design process we hypothesized we had created an ecologically oriented, comprehensive assessment instrument for third to fifth grade students that was developmentally appropriate and would gather developmentally valid information. Many researchers proceed from this stage to pilot testing the instrument with respondents, in order to collect data to examine the performance of the items and scales. This type of pilot testing can identify items that do not contribute to the statistical reliability of a scale, or that do not factor load with items intended to assess the same construct. These "bad items" can then be modified or eliminated. However, standard pilot testing does not provide information about *why* an item performed poorly. Cognitive pretesting is a methodology that can systematically provide such item performance information.

Cognitive Pretesting the ESSP-CF

Cognitive pretesting the ESSP-CF has been an iterative process for the ESSP project team; the methodology has evolved during its application to ESSP-CF items. At this time, two phases of cognitive pretesting have been completed, with modifications made to the ESSP-CF and the cognitive pretesting procedures after the first and second phases. Our efforts were aimed not only at improving the developmental validity of the ESSP-CF, but also at developing effective methods and procedures for future cognitive pretesting with children. The sample and methods used in each phase will be detailed below.

Sample and Methods: Phase 1—Teachers at an elementary school in north central North Carolina were trained to complete the Phase 1 cognitive pretesting procedure with 16 children. The purposive sample was chosen by the teachers to include 5 average readers and 11 below-average readers, all were African American third graders in an after school enrichment program. Reading status was based on teacher report. The gender of the children was not recorded; however, the sample included boys and girls. Teachers completed the cognitive pretesting procedure for 15 of the ESSP-CF items with each child respondent.

Phase 1 data were hand-recorded by the teachers on a recording sheet. Table 1 outlines the four-step cognitive pretesting procedure. For Step 1, teachers asked each child to read the item aloud and recorded any words children had trouble reading. In Step 2, teachers asked each child to put the item in his or her own words and recorded whether the child accurately interpreted the item, and if the child misinterpreted the item, how it was misinterpreted. In Step 3, teachers asked each child to choose an answer and recorded how long it took the child to choose. During Step 4, the teachers asked the child to explain the answer chosen and recorded what the child said in explanation of the answer choice, allowing the researchers to determine the validity of each response.

Sample and Methods: Phase 2—Phase 2 was conducted at an elementary school in western North Carolina, again employing teachers trained by the ESSP-CF research team to perform cognitive pretesting interviews. In the Phase 2 procedures, teachers audiotaped the cognitive pretesting interviews in order to preserve for analysis more complete data about child responses. This also allowed the teachers to more fully focus on the interview process. The researchers then coded the data directly from the tapes and transcripts of the tapes. The four-step interview process used in Phase 1 remained the same. The Phase 2 sample included 23 children, including 13 third graders and 10 fifth graders. This purposive sample was chosen by the teachers to include 11 girls and 12 boys, 11 European American and 14 African American students, 4 above-average readers, 11 average readers, and 8 below-average readers.

Analysis Procedures—The analysis during Phase 1 was based on the data on the recording sheets completed by the teachers. Problems with words, items, or answer options were summarized across children and across items. Once the data were summarized, and problem items identified, resultant modifications to ESSP-CF items were arrived at in a qualitative group process by two or three of the authors. In the analysis of the Phase 2 data, researchers listened to the tapes and completed similar but more detailed recording sheets than were completed by the teachers in Phase 1. The same qualitative group analysis was performed, and item modifications were made.

Findings—We anticipated that cognitive pretesting would form a valuable part of the design process, leading to improvements in the developmental validity of the ESSP-CF. The cognitive pretesting procedures were designed to assess three critical components of a developmentally valid item (Figure 1), including the ability of the child to read the item, to comprehend the item, and to give a valid explanation for the chosen response.

Our item development procedures prior to the cognitive pretesting, including input from five experts, resulted in items that most children could read without difficulty. In both phases of cognitive pretesting children were able to read the ESSP-CF items 96% of the time without assistance. However, cognitive pretesting revealed numerous problems with accurate interpretation of items and adequate response options, which resulted in significant modifications to the ESSP-CF. These changes included (a) redesigning one scale to make it less abstract, (b) changing answer option sets, and (c) rewording misinterpreted items.

Redesigning a scale: Phase 1 cognitive pretesting revealed the ESSP-CF self-esteem scale was too cognitively demanding. This scale included eight items such as “I am happy with myself,” “I feel good about myself,” and “I can think of some good things about myself.” Children in the targeted age range had difficulty with the level of abstraction of these items. For example, the item worded “I am happy with myself” was interpreted as a item about current mood; one child who answered this item “Sometimes” explained “Sometimes I feel happy, sometimes I don’t.” Piagetian theory about the concrete operations stage is consistent with this finding; the children were able to mentally represent the concepts in the items but

stumbled due to the abstract nature of a general assessment of self. We replaced the self-esteem scale with a more concrete set of items. Examples of the new items are: (a) I am smart, (b) I am good at art, (c) I am good at music, (c) I am proud of myself, and (d) I am good at sports. We postulated that this set of items would tap into a child's perception of self on a developmentally appropriate level, and indeed these items seemed to perform well in the Phase 2 cognitive pretesting.

Changing answer option sets: The Phase I version of the ESSP-CF included an option set consisting of the answers “Not like me,” “A little like me,” and “A lot like me.” This is a common answer set in self-report instruments for children; however, the cognitive pretesting revealed that some children found it confusing. We replaced the option set with one already used in the ESSP-CF: “Never,” “Sometimes,” “Often,” and “Always.” The original answer option set appears to have been abstract enough to cause the children difficulty. Although not as abstract as the self-esteem scale, the cognitive demands of evaluating something as “like me” are similar. Changing the option set also required some modifications to items. This change eliminated a source of measurement error and left fewer response sets in the child form (three, with two being very similar), which is a desirable outcome according to the scale development experts we consulted.

Rewording misinterpreted items: In the Phase 2 of cognitive pretesting it became apparent that children often missed the conditional context of an item, if the condition was stated at the end of the item. For example, the item “I can talk to grownups at my school when I need help” contains a conditional context at the end (*when* the child needs help). Some children chose the answer option “Sometimes,” explaining that sometimes they could get answers without help. Reversing the order of phrases to put the conditional statement first (“When I need help I can talk to grownups at my school”) sets the context of the item in a potentially more developmentally appropriate order. Ten items were reworded in this manner. The reordering of the concepts in these items will be assessed in future cognitive pretesting. Self-report items demand multiple parallel processes and may exceed the child's capacity when the order of the clauses require the child to find and encode the conditional statement *after* the core question.

Implications and Next Steps: The ESSP

Audiotaping the cognitive pretesting procedure allowed us to collect detailed information and reduce the demands on the teacher-pretesters, while not increasing the length of time of pretesting or the demands on the children. The tape-recorded data provided by child respondents while reading, interpreting and answering the ESSP-CF items were also preserved in a way that allowed a closer assessment of item performance. Additionally, the tapes provided an opportunity for assessment of the methodology itself. The most significant result of this analysis was modifying Step 2 (see Table 1) in our procedure. In this step, teachers were instructed to ask the children to “put the question in your own words.” Most children simply restated the item, or changed the order of the phrases or words in the item. This may have been partially due to the simple wording of the items, and possibly due to the somewhat abstract nature of the request. Interesting to note, this is a standard cognitive method procedure reported in the literature for use with adults (Forsyth & Lessler, 1991).

Knowing the goal was to assess the accuracy of interpretation of ESSP-CF items, the teachers apparently experimented with the procedure for Step 2. In Phase 2, four teachers used variations of “What does that question mean in your own words?” or “What does that question mean to you?” This approach did improve the quality of data collected to assess the accuracy of interpretation. A fifth teacher experimented with variations of: “What does the question mean, what is it asking?” Although the teacher used this technique for only three

items with one child, it did appear to be the most promising procedure. The child's responses were truly interpretive. For example, after the child read aloud the item "My friends help me when I am upset," the teacher asked "Could you tell us in your own words what that question is asking?" The child responded "Umm, do your friends care about you when you upset or feeling bad." We anticipate modifying our procedures for Step 2 in future cognitive pretesting in light of this finding.

We originally anticipated that children would be more comfortable during the cognitive pretesting procedure if the pretesting were conducted by familiar adults, such as teachers. While we still believe this is important, we now recommend that the researchers be present during at least each interviewer's initial pretesting interview. If present, the researcher can ensure procedural fidelity and provide guidance to interviewers about asking follow-up questions to obtain the necessary data from a child when the scripted questions are inadequate. Some researchers may choose to conduct the interviews themselves, an approach we plan to try in future applications of the methodology.

The design process of the ESSP is ongoing and has included numerous other procedures that are not discussed in this article (N. K. Bowen, Bowen, & Woolley, in press). Briefly, our future plans are to complete a third and fourth phase of cognitive pretesting. In Phases 3 and 4, we will assess the changes we made in Phase 2, and confirm the adequacy of the revised cognitive pretesting procedures. With the additional funding that the project team recently received from the National Institute on Drug Abuse, we will then conduct a large pilot test of the full instrument—child, parent/guardian, and teacher forms—at multiple schools. A one-site study is also planned that will use qualitative procedures to test how the ESSP gets implemented in schools and its impact on interventions and student outcomes.

Discussion and Applications to Social Work Practice

As a result of our experiences with applying and adapting cognitive pretesting to instrument design with children, we believe it is a promising qualitative method to advance the developmental validity of child self-report instruments. Cognitive pretesting as described in this article is a systematic approach to tapping the expertise of children about their own ability to successfully answer self-report items. From an argument-based approach to validity (Kane, 1992), cognitive pretesting strengthens the inference in the validity argument that children can read, comprehend and validly respond to the self-report items. Piagetian and information processing theories informed our conceptualization of the cognitive pretesting process.

Piagetian theory is helpful to guide the initial generation of self-report items for children, however, an information-processing model that stresses key components of reading, comprehension, and valid response explanation appears most useful in guiding the cognitive pretesting of self-report items for children. Assessing each component as a step in an information processing chain provides a systematic approach to pretesting the validity of an item presented to a child. This approach parallels the argument-based approach to establishing validity promoted by Kane (1992).

As social workers increasingly design and develop their own self-report measures for practice and research, adopting and adapting the most effective methods from the scale development and the survey design literatures will lead to higher quality measures. A systematic assessment of how clients interpret items seems necessary when developing instruments for any clients who may not interpret self-report items as designers intended due, for example, to age differences, cultural differences, language differences, and/or cognitive differences due to developmental delays, mental illness, or neurologically

degenerative disorders. Cognitive pretesting is one promising method for conducting this type of systematic assessment.

We look forward to seeing other social work researchers apply cognitive methods including cognitive pretesting in the development of self-report instruments. We anticipate that other researchers will find this qualitative method to advance the validity of self-report instruments helpful when designing instruments for children and other client populations who are the target of social work services.

References

- Alwin DF. Research on survey quality. *Sociological Methods and Research*. 2001; 20:3–29.
- Amaya-Jackson L, Socolar RRS, Hunter W, Runyan DK, Colindres R. Directly questioning children and adolescents about maltreatment: A review of the survey measures used. *Journal of Interpersonal Violence*. 2000; 15:725–759.
- Bowen, GL.; Richman, JM. The School Success Profile. Chapel Hill: School of Social Work University of North Carolina; 1995, 1997, 2001.
- Bowen GL, Woolley ME, Richman JM, Bowen NK. Brief intervention in schools: The School Success Profile. *Brief Treatment and Crisis Intervention*. 2001; 1:43–54.
- Bowen, NK.; Bowen, GL.; Woolley, ME. Constructing and validating assessment tools for school-based practitioners: The Elementary School Success Profile. In: Roberts, AR.; Yeager, K., editors. *Handbook of Practice Based Research*. New York: Oxford University Press; in press
- Cooley MR, Turner SM, Beidel DC. Assessing community violence: The children's report of exposure to violence. *Journal of the Academy of Child and Adolescent Psychiatry*. 1995; 34:201–208.
- DeMaio, TJ.; Rothgeb, JM. Cognitive interviewing techniques: In the lab and in the field. In: Schwarz, N.; Sudman, S., editors. *Answering Questions: Methodology for Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass; 1996. p. 177-196.
- DeVellis, RF. Scale development: Theory and applications. Thousand Oaks, CA: Sage; 1991.
- Ferguson TJ, Stegge H, Eyre HL, Vollmer R, Ashbaker M. Context effects and the (mal)adaptive nature of guilt and shame in children. *Genetic, Social, and General Psychology Monographs*. 2000; 126:319–345.
- Foddy W. An empirical evaluation of in-depth probes used to pretest survey questions. *Sociological Method and Research*. 1998; 27:103–133.
- Forsyth, BH.; Lessler, JT. Cognitive laboratory methods: A taxonomy. In: Biemer, PP.; Groves, RM.; Lyberg, LE.; Mathiowietz, NA.; Sudman, S., editors. *Measurement Errors in Surveys*. New York: Wiley & Sons; 1991. p. 393-418.
- Fowler, FJ. *Improving Survey Questions*. Thousand Oaks, CA: Sage; 1995.
- Gambrill E. Evidence-based practice: An alternative to authority based practice. *Families in Society: The Journal of Contemporary Human Services*. 1999; 80:341–350.
- Goldhaber, DE. *Theories of Human Development: Integrative perspectives*. Mountain View, CA: Mayfield; 2000.
- Goswami, U., editor. *Blackwell handbook of childhood cognitive development*. Malden, MA: Blackwell; 2002.
- Greenwald R, Rubin A. Assessment of posttraumatic symptoms in children: Developmental and preliminary validation of parent and child scales. *Research on Social Work Practice*. 1999; 9:61–75.
- Hastie, R. Information processing theory for survey researchers. In: Hippler, HJ.; Schwartz, N.; Sudman, S., editors. *Social information processing and survey methodology*. New York: Springer-Verlag; 1987.
- Jabine, TB.; Straf, ML.; Tanur, JM.; Tourangeau, R. Cognitive aspects of survey methodology: Building a bridge between disciplines. Washington DC: National Academy Press; 1984.
- Jobe JB, Mingay DJ. Cognitive research improves questionnaires. *American Journal of Public Health*. 1989; 79:1053–1055. [PubMed: 2751028]

- Kane MT. An argument-based approach to validity. *Psychological Bulletin*. 1992; 112:527–535.
- March JS, Parker JDA, Sullivan K, Stallings P, Conners CK. The multidimensional anxiety scale for children (MASC): Factor structure, reliability, and validity. *Journal of the Academy of Child and Adolescent Psychiatry*. 1997; 36:554–565.
- McConaughy SH. Evaluating behavioral and emotional disorders with the CBCL, TRF, and YSR cross-informant scales. *Journal of Emotional and Behavioral Disorders*. 1993; 1:40–52.
- McGuire S, Manke B, Saudino KJ, Reiss D, Hetherington EM, Plomin R. Perceived competence and self-worth during adolescence: A longitudinal behavioral genetic study. *Child Development*. 1999; 70:1283–1296. [PubMed: 10621957]
- Montgomery MS. Self-concept and children with learning disabilities: Observer-child concordance across six context-dependent domains. *Journal of Learning Disabilities*. 1994; 27:254–262. [PubMed: 8051506]
- Rebok G, Riley A, Forrest C, Starfield B, Green B, Robertson J, et al. Elementary school-aged children's reports of their health: A cognitive interviewing study. *Quality of Life Research*. 2001; 10:59–70. [PubMed: 11508476]
- Schwartz N. Self-reports. *American Psychologist*. 1999; 54:93–105.
- Sudman, S.; Bradburn, NM.; Schwartz, N. *Thinking about answers*. San Francisco: Jossey-Bass; 1996.
- Tanur, JM. Looking backwards and forwards at the CASM movement. In: Sirken, MG.; Herrmann, DJ.; Schechter, S.; Schwarz, N.; Tanur, JM.; Tourangeau, R., editors. *Cognition and survey research*. New York, NY: Wiley-Interscience; 1999. p. 13-20.
- Thomas EJ. The validity of design and development and related concepts in developmental research. *Social Work Research & Abstracts*. 1985; 21:50–55.
- Thyer, BA. *The handbook of social work research methods*. Thousand Oaks, CA: Sage; 2001.
- Tinsley BJ, Holtgrave DR. A multimethod analysis of risk perceptions and health behaviors in children. *Educational & Psychological Measurement*. 1997; 57:197–209.
- Tourangeau, R.; Rips, LJ.; Rasinski, K. *The psychology of survey response*. Cambridge, UK: Cambridge University Press; 2000.
- Willis, GB.; DeMaio, TJ.; Harris-Kojetin, B. Is the bandwagon headed for the Promised Land? Evaluating the validity of cognitive interviewing techniques. In: Sirken, MG.; Herrmann, DJ.; Schechter, S.; Schwarz, N.; Tanur, JM.; Tourangeau, R., editors. *Cognition and survey research*. New York, NY: Wiley-Interscience; 1999. p. 133-154.
- Wright-Strawdermann C, Watson BL. The prevalence of depressive symptoms in children with disabilities. *Journal of Learning Disabilities*. 1992; 25:258–264. [PubMed: 1573336]
- Yugar JM, Shapiro ES. Elementary children's school friendship: A comparison of peer assessment methodologies. *School Psychology Review*. 2001; 30:568–587.

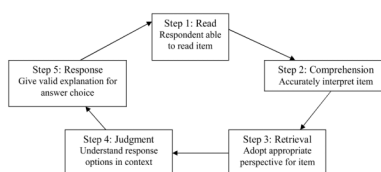


Figure 1.
Cognitive Pretesting Model for Children

Note: Informed by DeMaio & Rothgeb (1996), Jobe & Mingay (1989), and Tourangeau, Rips, & Rasinski, (2000).

Table 1**ESSP-CF Cognitive Pretesting Steps**

Step 1	Ask the child to read aloud the question on the screen.
Step 2	Ask the child to paraphrase the question or put the question into his/her own words.
Step 3	Ask the child to pick the best answer to the question.
Step 4	Ask the child to explain his/her answer.

Note: Phase 1—a recording sheet was completed by the teachers for each item pretested. Additional data were collected at each step such as words the child had difficulty reading, accuracy of item interpretation, elapsed time to report response, and what child said to explain chosen answer. Phase 2—cognitive pretesting interviews were audiotaped, and researchers completed recording sheets from the audio tapes.